DOCUMENT RESUME

ED 291 801 TM 011 113

AUTHOR Puleo, Vincent T.; Lieberman, Billie E.
TITLE Maximizing the Use of Standardized Test Results: A Brief Case Study.
PUB DATE [86]
NOTE 21p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Records; Data Analysis; Elementary Secondary Education; Item Analysis; *Remedial Instruction; *Standardized Tests; Testing Programs; Test Interpretation; *Test Results; Test Use
IDENTIFIERS New York (Kingston)

ABSTRACT

This preliminary investigation explored a means of maximizing the use of standardized test results. In addition to the usual package of alpha and ranked norm-referenced test reports, a special series of reports by the California Test Bureau/McGraw-Hill was generated for remedial reading and mathematics teachers. These reports displayed primarily objectives or skill-referenced information. To assess staff reaction, a questionnaire was completed by the remedial teachers, and interviews were conducted with administrators involved. Both teacher and administrative reactions to the special reports were encouraging. The experience indicates that it is possible to obtain greater use of existing standardized test results through the use of special reports. In relation to earlier studies that question the relevance of standardized tests, current findings also prompt consideration of whether the special reports would favorably influence classroom teacher perception; and whether there are differences in the perception of standardized tests between classroom and remedial teachers. (The study instrument is appended.) (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED291801

Maximizing The Use Of Standardized Test Results: A Brief Case Study

Vincent T. Puleo and Billie E. Lieberman

Kingston City Schools, Consolidated

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

VINCENT T. PULEO

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 011 113

Running Head: STANDARDIZED TESTING

BEST COPY AVAILABLE

Abstract

This preliminary investigation explored a means of maximizing the use of standardized test results. In addition to the usual package of alpha and ranked norm-referenced test reports, a special series of reports by the California Test Bureau/McGraw-Hill were generated for remedial reading and mathematics teachers. These displayed primarily objectives or skill-referenced information.

To assess staff reaction, a questionnaire was completed by the remedial teachers and interviews were conducted with administrators involved. Both teacher and administrative reaction to the special reports were encouraging. The experience indicates it is possible to obtain greater use of existing standardized test results through the use of special reports.

In relation to earlier studies which question the relevance of standardized tests, current findings also prompt consideration of whether the special reports would favorably influence classroom teacher perception; and whether there are differences in the perception of standardized tests between classroom and remedial teachers.

Maximizing The Use Of Standardized Test Results: A Brief Case Study

This report describes preliminary study of an attempt to maximize the use of district-wide standardized testing results. The premise for the study is that, since we are required by Federal Chapter I regulations to test our students each spring for remedial reading and mathematics screening, the data derived from testing needed to be explored for other possible valuable uses. What we found were some encouraging results, and intriguing questions for further study.

Background

In Kingston, New York we probably make more use of our standardized testing results than most districts. We routinely get alphabetic and ranked student lists, and frequency distributions by building and district. We get item analyses selectively, and on occasion we have had special grouping reports, e.g., out-going 6th grade students re-arranged by feeder junior-high school. These have all been grade-wide and building reports. In addition, we do a variety of in-house analyses. But our experience thus far has been limited to the standard survey or norm-referenced interpretation of these

results. We have pursued this approach while remaining aware that our testing inherently contains a great deal more information.

We wanted to use the results of our standardized testing in keeping with a diagnostic-prescriptive model of instruction (Stellern, Vasa, & Little, 1976). At the same time we kept in mind research which pointedly illustrates difficulties in the relevance of standardized testing as well as that which gives insight into classroom decision-making and assessment.

Research on teacher concerns with standardized testing has particularly included problems with its timing, e.g., solely at the beginning or end of the school year, the global nature of portraying results, e.g., grade-equivalent scores or percentile ranks, and with the practice of referencing pupil performance to an external group rather than to instructional objectives (e.g., Goslin, 1967; Salmon-Cox, 1981; Shulman, 1980). Insights provided by the literature indicate that appropiate testing can enhance the focus of instruction, and strengthen communication between teacher and student (Gullickson, 1984). In addition, there is clearly an interrelatedness between various forms of assessment; that is, different tests serve different functions for pupils, teachers, and administrators (Puleo, 1986). Thus, for remedial teachers we wanted objectives or

skill-referenced information primarily, and we wanted this in their hands in early September. For administrators we wanted the same data aggregated by building and grade. And, we wanted to do so while avoiding one of the premier hazards of standardized testing results - data overkill; that is, an inordinate quantity of superfluous data.

As Kingston uses the Comprehensive Tests of Basic Skills (CTBS) (California Test Bureau/McGraw-Hill, 1981), we discussed this challenge with our California Test Bureau/McGraw-Hill (CTB) representative. A set of special computer reports were presented for consideration. This CTB report service has been available for some years. However, to our knowledge there has been no systematic study of its effects on student, teacher or administrative performance.

The following is an abbreviated accounting of our experience with this project during the 1985-86 school year. The test results were from the spring of 1985.

Method

Identification of our Report Population

In screening for remedial candidates, we want to insure that all students subsequently selected would be included in our special reports. Therefore, the decision was made to use the national norm with some

additional margin for measurement error. Pupil performance at or below the 55th percentile was identified as the criterion for inclusion in the reports.

This criterion was applied to the CTBS Vocabulary, Reading Comprehension and Total Math subtests for grades 1 to 12. It resulted in approximately 4000 students being included in the report series. The students were drawn from ten elementary schools, three junior highs, and one senior high school.

Criteria For Report Selection

To guide report selection we established the following criteria:

1. The reports should identify our remedial population. By Federal Chapter I guidelines we had to use the 23rd percentile. However, since we want to insure that all needy students receive services, we use the local 23rd percentile as opposed to the national one. This is a more rigorous standard.

2. The reports should provide us with a skill profile of the child at-risk so we could pinpoint instructional areas of concentration. This would be the basis of our diagnostic and prescriptive approach.

3. The reports should identify skill deficiency trends in individual schools, and district by grade.

Reports Selected

We reviewed available CTB report formats and chose those best fitting our criteria.

Three reports were identified for teacher use. The Individual Test Record (ITR) lists each objective within the identified subject areas, and portrays non-mastery, partial knowledge or mastery of each. It also shows the precision by which an item was mastered and the number of items not reached by the student. This gives the teacher a valuable overview of strengths and weaknesses.

The ITR shows test results in terms of Scaled Scores, Grade Equivalent, Normal Curve Equivalent, National Stanines, Local Percentiles and National Percentiles. A graphic portrayal is also shown of the individual's test performance based on national percentiles.

The Right Response Record (RRR) is a report for each grade in a particular school. It lists all the students at each grade in that school and tells item by item if the student passed, failed, or left the item blank. This report shows how a student performed in relation to the group. It also can help a diagnostician pick out particular students who need specific skill help. The RRR further indicates how the group did on each skill tested. For example, it shows the group average in relation to the total number of items in that

objective. At the same time it reveals how the standardization population performed on that objective.

The third report is the Selected Student Report (SSR). This is a summary listing by subject area of those students who were identified according to the cut-off point specified, e.g., 55th percentile. This produces a convenient student class list.

For administrative use, two additional reports were selected. These were building and grade summaries of the RRR.

Teacher Use

Reports were generated for all 26 remedial staff. Overall, these teachers were highly experienced. However, some had no formal training in the teaching of reading to remedial students.

Because of scheduling delays and minor difficulties with formatting, reports were distributed to staff in mid-October. At that time a training session was held and each report was reviewed in detail. Teachers were instructed in techniques of how to identify the skill needs of each student.

Local diagnosis and prescription forms containing information derived from the reports were also completed at this time. These completed forms became an integral part of the instructional plan for the remedial population.

Project Evaluation

The effects of the special reports were evaluated in this preliminary investigation in two ways: a survey of remedial teachers, and interviews with administrators involved. Both are viewed as formative in nature.

Results

Teacher Survey Results

Twenty-five of the 26 remedial teachers, or 96.2%, completed and returned a questionnaire designed to elicit reactions to the report series. The questionnaire is shown in Appendix A. Nine aspects of the reports were assessed. Table 1 shows the results.

Insert Table 1 about here

Observationally, this indicates general support for the project. On 7 of the 9 points, for example, a minimum of 60% of the teachers agreed or strongly agreed that there was value for them. The item gaining least support was the national and local comparison by objective on the RRR. Only 28% felt this had value while a substantial percent, 32, had no opinion. In retrospect, reaction to this norm-referenced point, in

light of the diagnostic-prescriptive thrust of the project, might have been predicted. The other point which reflects relatively less support from this teacher sample pertains to the top graphic display on the ITR. The survey indicated 56% felt this was useful. Also, a relatively large percent, 21, did not care to express an opinion. Interestingly, the graphic display is another piece of norm-referenced information. The last point, dealing with teacher desirabitity of having the reports again next school year, was informative. Here, 68% were affirmative. Considering the delay in project implementation, that is, mid-October, and its pilot or developmental status this first year, these results are viewed as encouraging.

Teacher comments richly supplemented the questionnaire data. Recurring thoughts included the potential value of having similar information for classroom teachers, the overlap in information between the reports and recording devices in use as part of the basal reading series, and the greatly increased quantity of material teachers are given. Additional in-service was indicated as desirable and, not unexpectedly, the importance of having the reports available early in September was stressed. A question was also raised about the value of the reports in relation to cost.

As the questionnaire did not elicit information

regarding how the reports contributed to teacher instruction, follow-up interviews were conducted on a non-systematic basis. This tended to indicate that the reports were helpful in a manner congruent with the literature. That is, clarity and focus in instruction were improved, and instructional communication with students was enhanced. This is obviously an area subsequent investigation should pursue.

Administrative Use

Administrative reaction remains in narrative form due to the small number of individuals involved.

We learned that the reports were of great value to administrators. In effect, they established a management information system whereby program or curriculum effects could be monitored. For instance, the director of remedial services was able to pinpoint performance on different levels of aggregation. One particular event illustrates this point. An elementary school in the district scored significantly below the others. And the second grade in the school was pointedly lower than the second grades in the rest of the district. From the summary reports, the director was able to determine in which classroom the problem was greatest, what specific skill needs were lacking, and what items under each of those skill areas were failed. Having this information available was essential because

initial teacher reaction did not view the matter as amenable to instructional correction.

Having made the teachers aware of the problem, a plan was devised to improve student performance in that building. Teachers were shown the skill deficiences in greatest need of attention. The spring 1986 test results showed that we should pursue the needs of these students further.

In addition to remedial staff, faculty meetings with classroom teachers in select schools were devoted to consideration of the special reports. The value of the reports was illustrated by going over the school's norm-referenced test results and explaining how the special information could be useful. The greater instructional specificity of the reports was clearly revealed.

Discussion

This preliminary study explored a means of maximizing the use of standardized test results. The effects were sufficiently supportive to continue, and possibly expand, the report series next year. At the same time, we learned of ways to strengthen the project.

Foremost was an awareness that remedial staff needed additional training in the use of the report data. We also need to schedule time so that classroom

and remedial teachers could share the information and plan coordinated instruction. Our experience likewise suggests that classroom teachers and building principals would benefit from a similar series of reports.

On another level of inquiry, we learned that it is essential to gain a greater understanding of the specific ways in which the report data complement existing sources of information, e.g., from the basal reading program. In particular, study is needed to identify what instructional functions in the various phases of the annual instructional cycle are served by this information. It might be proposed that the CTB reports serve as a CRT monitoring device (Puleo, 1986), and hence accomplish an important function within an information system.

Implications Regarding Teacher Perception of Standardized Tests

Current findings are provocative in light of earlier research which questions the relevance of standardized tests for classroom teachers (e.g., Goslin, 1967; Salmon-Cox, 1981; Shuman, 1980). More pointedly, two related questions are raised:

1. Might classroom teacher perception be more favorable if they received the special reports?

2. Is there a difference in the perception of standardized tests between remedial and classroom

teachers?

Future Use Of The CTB Special Reports

For the 1986-87 school year a number of changes will be made. The SSR will be dropped as staff perceives this as having relatively little value. An additional CTB service we will pilot in three elementary schools is a reshuffling of our spring test results into fall class lists. This would place at the fall teacher's fingertips all the results of our most recent testing cycle. With reshuffling the principal can also have the scope of student abilities at hand. This should be of distinct value in curriculum review and planning, as well as in classroom visitations or observations.

Consideration is likewise being given to reproducing Individual Program Reports for parents. While this is not a new idea, we feel that, within a diagnostic-prescriptive orientation, a portrayal of pupil performance by learning objectives in addition to the standard norm-referenced interpretation will strengthen parent-teacher communication. These reports would be used during conference time when they would be fully explained to parents. Finally, we're going to lower our cut-off score for student identification to the 50th national percentile. This will yield a degree of economy, and result in smaller, more manageable lists

for teachers to process for remedial eligibility.

Future Program Evaluation

Subsequent program evaluation will move to a summative approach. This will examine the effect of the report series as determined by the pattern and degree of standardized test differences before and after the time of implementation.

Summary and Conclusions

This preliminary investigation studied an attempt to maximize the use of standardized test results. In addition to the usual package of alpha and ranked norm-referenced results, a special series of reports by the California Test Bureau/McGraw-Hill were generated for remedial teachers for our spring 1985 Comprehensive Tests of Basic Skills. These displayed primarily objectives or skill-referenced information.

A number of initial procedural difficulties, including minor formatting changes, delayed return of the reports until mid-October. Staff training in their use was undertaken at this time with assistance of the CTB representative.

To assess staff reaction, a questionnaire was completed by the remedial teachers and interviews were conducted with administrators involved. Both teacher and administrative reaction to the special reports were

encouraging. These results indicate it is possible to make greater use of existing standardized test results.

In relation to earlier studies which question the relevance of standardized tests, current findings also prompt consideration of whether the special reports would favorably influence classroom teacher perception; and whether there are differences in the perception of standardized tests between classroom and remedial teachers.

Plans were described which would refine and extend our use of the special reports. Future project evaluation includes study of the effects of the reports on pupil test performance. Also needed is study of the specific ways the reports serve teacher instruction.

References

California Test Bureau/McGraw-Hill. (1981). Comprehensive Tests of Basic Skills, Form U. Monterey, CA: Author.

Goslin, D. A. (1967). Teachers and testing. New York: Russell Sage Foundation.

Gullickson, A. R. (1984). Teacher perspectives on their instructional use of tests. Journal of Educational Research, 77, 244-248.

Puleo, V. T. (1986). Anatomy of a testing program. Journal of the School Administrators Association of New York State, 17, 21-24.

Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? Phi Delta Kappan, 62, 631-634.

Shulman, L. S. (1980). Test design: A view from practice. In E. L. Baker & E. S. Quellmalz (Eds.), Educational testing and evaluation: Design, analysis and policy. Los Angeles, CA: Sage.

Stellern, J., Vasa, S. F., & Little, J. (1976). Introduction to diagnostic-prescriptive teaching & programming. Glen Ridge, NJ: Exceptional Press.

Appendix A

June 16, 1986

TO: Remedial Staff

Please help us out.

In connection with the CTBS results, this year we used a special set of CTB reports in addition to the usual BOCES test reports.

We need your reaction to these special reports so that we can plan for next year.

Respond to the following statements. You may want to refer directly to the reports in completing this form. If something important is not addressed by these questions, indicate your comments/concerns in the Comments Section.

Thank you--

Signed
Director of Reading and Library Media Services

	Strong Agree	Agree	No Opin.	Disag.	Strong Disag.
Regarding the Individual Test Record:					
1. The display at the top showing converted scores, e.g., NCE, NS, was useful to me	_____	_____	_____	_____	_____
2. The graphic display at the top to the right of the converted scores showing bands of error was useful to me. . . .	_____	_____	_____	_____	_____
3. The display at the bottom showing objective mastery scores according to Non-Mastery, Partial Mastery or Mastery was useful to me	_____	_____	_____	_____	_____
4. In general, I found the Individual Test Record useful to me as a guide for instruction	_____	_____	_____	_____	_____
Regarding the Right Response Record:					
5. The display showing pass/fail for each item for each objective was useful to me	_____	_____	_____	_____	_____

	Strong Agree	Agree	No Opin.	Disag.	Strong Disag.
6. The comparison between the items right nationally and locally was useful to me	_____	_____	_____	_____	_____
7. In general, I found the Right Response Record useful to me as a guide to instruction	_____	_____	_____	_____	_____
Regarding the Selected Student Report:					
8. I found this report useful to me . . .	_____	_____	_____	_____	_____
9. I would like to have all of these reports available next year	_____	_____	_____	_____	_____

Comments

Table 1
Remedial Teachers Responses to
Special Report Series (Percentages)

	Strong Agree	Agree	No Opin.	Disag.	Strong Disag.
Individual Test Record					
Top Display Showing Converted Scores	36	52	0	12	0
Top Graphic Display	4	52	20	20	4
Mastery Display	24	48	12	12	4
General Usefulness	20	52	8	16	4
Right Response Record					
Pass/Fail Display	4	64	4	24	4
Nat'l/Local Comparison	0	28	32	36	4
General Usefulness	8	52	12	24	4
Selected Student Report					
General usefulness	8	68	4	16	4
Desirability of having reports next year	16	52	16	12	4